

学校编码: 10384

分类号_____密级_____

学号: 21720061152215

UDC _____

厦 门 大 学

硕 士 学 位 论 文

HIV 包膜蛋白在酵母系统中的表达、纯化及其应用研究

**Expression and Purification of HIV-1 Envelope
Glycoprotein gp160 and gp120 in Yeast and Its Study**

何 芳 萍

指导教师姓名: 夏宁邵教授

专 业 名 称: 生物化学与分子生物学

论文提交日期: 2009 年 7 月 8 日

论文答辩时间: 2009 年 8 月 3 日

2009 年 8 月

厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下,独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果,均在文中以适当方式明确标明,并符合法律规范和《厦门大学研究生学术活动规范(试行)》。

另外,该学位论文为(夏宁邵 教授)课题(组)的研究成果,获得(夏宁邵 教授)课题(组)经费或实验室的资助,在(夏宁邵 教授)实验室完成。(请在以上括号内填写课题或课题组负责人或实验室名称,未有此项声明内容的,可以不作特别声明。)

声明人(签名):

年 月 日

厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

（ ） 1. 经厦门大学保密委员会审查核定的保密学位论文，
于 年 月 日解密，解密后适用上述授权。

（ ） 2. 不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年 月 日

摘 要

获得性免疫缺陷综合征（AIDS），是由人类免疫缺陷病毒（human immunodeficiency virus, HIV）感染所引起的一种严重的传染性疾病。由于 HIV 的高度变异，尚缺乏有效的预防性疫苗，目前血清学诊断和社会防控是控制 HIV 传播的主要手段。研究表明，HIV 的中和抗体和特异血清标志主要位于包膜蛋白 Env 上，对于 HIV 包膜蛋白的研究主要是通过多种表达系统进行的。本研究利用酿酒酵母和毕赤酵母两种表达系统对 HIV 膜蛋白进行表达、纯化和性质研究。

利用生物信息学软件 GCG 对 HIV gp160 蛋白进行二级结构预测和疏水性及潜在糖基化位点分析，结果表明 aa511-aa537, aa679-aa703 肽段是 gp160 中最强的疏水区域，可能影响蛋白的可溶性；构建了强疏水区段缺失的 gp160 突变体，在酿酒酵母中表达，获得可溶性较好的 rgp160 Δ 12 蛋白，且呈现为糖基化的三聚体形式，对于 HIV 阳性血清尤其是早期感染的血清具有良好的反应性，提示其可能应用于 HIV 各类诊断试剂的价值。

利用毕赤酵母分泌表达 gp120 及其 N 端缺失突变体，结果表明，N 端缺失 60aa 的表达量最高，而 N 端缺失 90aa 则不能表达，提示 N 端氨基酸的分布情况对于 gp120 的表达和/或分泌途径有显著的影响；分泌表达的 rgp120 较大肠杆菌来源的 gp120，对 HIV 阳性血清有较高的检出率，提示其可能较好地模拟了 gp120 的天然表位构象。

利用多尺度的发酵条件摸索，获得了酿酒酵母胞内表达和毕赤酵母分泌表达的高密度发酵工艺，经过纯化方法的优化，rgp160 Δ 12 和 rgp120 蛋白的最终有效表达量分别约为 10mg/L 和 30mg/L，从而为 HIV 诊断试剂和疫苗研制奠定了良好的基础。

关键词：人免疫缺陷病毒（HIV） 包膜蛋白（Env） 酿酒酵母 毕赤酵母

Abstract

Worldwide, AIDS is a very dangerous disease and poses serious harm to health. This makes the study of HIV/AIDS therapy and prevention extremely important. But given the high variation in the virus, there is no safe vaccine or effective anti-HIV therapy. Control of the HIV epidemic depends mainly on early detection of the infected. HIV infection or immunization with Env-based immunogens induces the production of antibodies which can be divided into six major classes based on the location and properties of their epitopes. Our objective was to express the envelope glycoproteins gp160 and gp120 from the virus pNL4-3 in *Saccharomyces cerevisiae* and *Pichia pastoris*.

Two strong hydrophobic regions, aa 511-537 and aa 679-703, were predicted by GCG Wisconsin Package software and removed to investigate the solubility of the mutated gp160 (gp160 Δ 12). The results showed that rgp160 Δ 12 was soluble enough as to be present exclusively in the supernatant of cell lysate. The mutant exists as a trimer in solution via intermolecular disulfide bonds. In addition, rgp160 Δ 12 carries distinct epitopes which react with early-HIV-infection antibodies, which may provide a novel antigen for immunodetection assays.

The rgp120 was expressed as secretion in *Pichia pastoris*. A 60aa N-terminus deletion improved expression levels, while a 90aa deletion lowered them. ELISA performed using a panel of human sera suggests that purified rgp120 had better biologic activity than gp120 which was expressed in *E. coli*, namely that rgp120 might better emulate the native conformation of gp120.

Fermentation procedure was optimized for high cell-density yield. Rgp160 Δ 12 expression level was about 10mg/L and rgp120 about 30mg/L. These two may provide novel antigens for immunodetection assays, vaccine candidates or other related research purposes.

Key words: HIV; envelope glycoproteins; *Saccharomyces cerevisiae*; *Pichia pastoris*

缩略词

缩写	英文全称	中文名称
°C	degree celsius	摄氏度
μg	microgram	微克
μL	microlitre	微升
aa	amino acid	氨基酸
AIDS	acquired immunodeficiency syndrome	获得性免疫缺陷综合症
Amp	Ampicillin	氨苄青霉素
AOX	Alcohol Oxidase	醇氧化酶
Ap	Alkaline phosphatase	碱性磷酸酶
bp	base pair	碱基对
BSA	bovine serum albumin	牛血清蛋白
CD4	cluster of differentiation 4	
CO ₂	carbon dioxide	二氧化碳
Da	Dalton	道尔顿
DNA	Deoxyribonucleic Acid	脱氧核糖核酸
DO	Dissolved Oxygen	溶解氧
ELISA	Enzyme-linked ImmunoSorbant Assay	酶联免疫吸附测定
Env	Envelope protein	包膜蛋白
g	gravity	重力速度
h	hour	小时
HA	Hemagglutinin	血凝素
HIV	Human immunodeficiency virus	人免疫缺陷病毒
HRP	Horseradish Peroxidase	辣根过氧化物酶
Kan	Kanamycin	卡那霉素
kb	kilo base pair	千碱基对
kD	kilo Daltons	千道尔顿

L	Litre	升
M	Mol/L	摩尔/升
Man	mannose	甘露糖
min	minute	分钟
Mut	Methanol utilization phenotype	甲醇利用表型
MW	Molecular Weight	分子量
PCR	polymerase chain reaction	聚合酶链式反应
rpm	revolutions per minute	每分钟转速
RT	reverse transcriptase	反转录酶
s	second	秒
SIV	Simian immunodeficiency virus	猿免疫缺陷病毒
SU	surface glycoprotein	表面膜蛋白
TM	transmembrane glycoprotein	跨膜蛋白
UNAIDS	the Joint United Nations Programme on HIV/AIDS	联合国艾滋病规划署
WB	Western Blot	蛋白印迹实验
WHO	World Health Organisation	世界卫生组织

注：按缩略词首字母排序

摘 要.....	I
ABSTRACT.....	II
缩略词.....	III
前 言.....	1
1: 人类免疫缺陷病毒.....	1
1.1 HIV 概况.....	1
1.2 HIV 的分子生物学特征.....	2
1.3 HIV 包膜糖蛋白的研究进展.....	5
1.4 HIV 感染的实验室诊断研究进展.....	14
1.5 HIV 的疫苗研究.....	17
2: 酵母表达系统.....	18
2.1 酵母表达系统概况.....	18
2.2 酿酒酵母表达系统.....	20
2.3 巴斯德毕赤酵母 (<i>Pichia pastoris</i>) 表达系统.....	22
2.4 酵母糖基化修饰能力.....	26
3: 本文研究的目的是与意义.....	28
材 料 与 方 法.....	30
1: 材料.....	30
1.1 主要仪器.....	30
1.2 主要试剂和材料.....	31
1.3 常用溶液和培养基.....	32
2: 方法.....	39
2.1 分子克隆.....	39
2.2 酵母相关实验.....	42
2.3 蛋白质性质研究方法.....	45
2.4 计算机辅助分析与设计.....	48
结 果 与 分 析.....	49
1. HIV 包膜蛋白的性质分析.....	49
2. rgp140、rgp160 及其突变体 rgp160 Δ 12 在酿酒酵母中的表达与性质鉴定.....	50
2.1 重组表达菌株的构建.....	50
2.2 rgp160、rgp140 和 rgp160 Δ 12 的表达.....	53
2.3 rgp160 Δ 12 发酵工艺的摸索.....	54
2.4 rgp160 Δ 12 的纯化工工艺摸索.....	55

2.5 rgp160 Δ 12 蛋白特性的分析.....	56
3. rgp120 在毕赤酵母中的表达与性质鉴定	58
3.1 重组表达菌株的构建	58
3.2 酵母电转化和转化子的筛选.....	60
3.3 重组蛋白 rgp120 的诱导表达	61
3.4 重组蛋白 rgp120 的表达摸索	63
3.5 重组蛋白 rgp120 的纯化与性质研究	64
4. rgp160、rgp160Δ12、rgp120、rgp120N60 对艾滋病检测参比血清的反应结果.....	66
4.1 酿酒表达的 rgp160 Δ 12 的参比血清反应结果	66
4.2 rgp120、rgp120N60 的参比血清反应结果.....	67
4.3 rgp120 的 Western Blot 活性分析.....	68
讨 论.....	70
1. HIV 包膜蛋白与表达系统的选择.....	70
2. 重组包膜蛋白的表达摸索与发酵放大工艺摸索	71
2.1 重组包膜蛋白的表达摸索.....	71
2.2 发酵放大工艺摸索.....	73
3. 重组包膜蛋白的纯化与性质鉴定.....	74
3.1 rgp160、rgp120 的纯化.....	74
3.2 重组胞膜蛋白的性质鉴定	75
小 结.....	76
展 望.....	77
参 考 文 献.....	78
致谢.....	87
附录.....	88

前言

1: 人类免疫缺陷病毒

1.1 HIV 概况

人类免疫缺陷病毒（human immunodeficiency virus, HIV）是导致艾滋病（AIDS）这一全球性问题的致病因子[1]。HIV 感染者终生携带病毒，多数感染者在数年的潜伏期后发病，最终几乎都将死于反复感染或肿瘤。自 1981 年发现 AIDS 以来[2]，AIDS 病在世界范围内迅速蔓延，感染者终生携带病毒。据 WHO/UNAIDS 估计，截至 2007 年底，全世界因艾滋病死亡的人数累计达到了 2,000 多万，3,300 多万人为感染者，其中撒哈拉南部非洲是现在 AIDS 流行最为严重的地区，全球 60% HIV 感染者生活在这一地区（图 1-1）[3]。我国的 HIV 感染者虽然只占世界的 0.5% 左右，但 HIV/AIDS 已从局部流行进入到广泛流行的长期增长期，为全球增长最快的地区之一，就 2007 年，新发感染者 5 万，因 AIDS 死亡 2 万[4]。

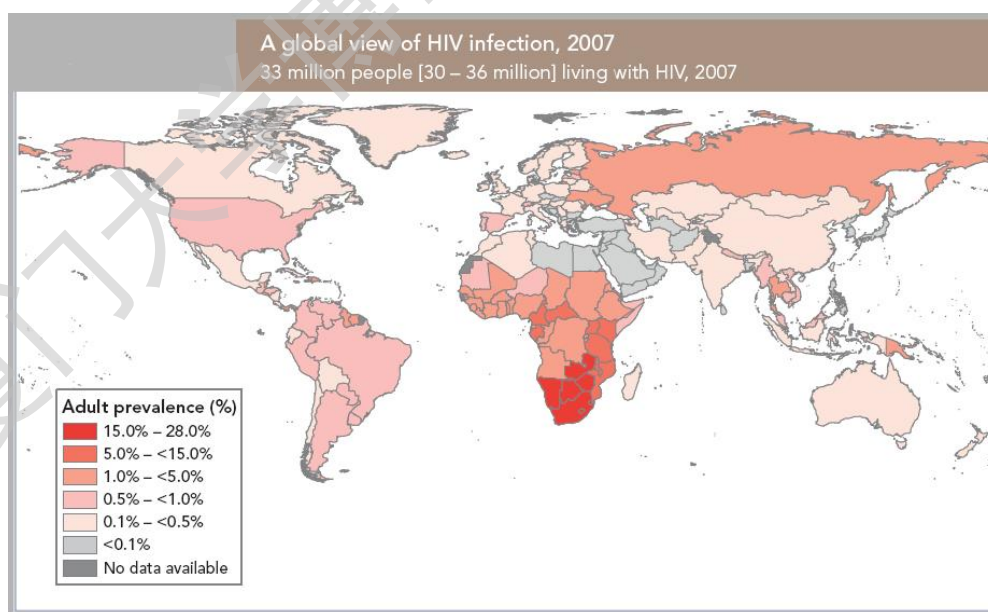


图 1-1 2007 年全球 HIV/AIDS 感染者分布情况

Fig. 1-1 A global view of HIV infection, 2007

图片来源:UNAIDS report 2008

艾滋病病毒在分类学上属于反转录病毒科慢病毒属中的灵长类慢病毒群[5]。迄今为止，全球流行的 HIV，根据血清学反应和病毒核酸序列测定可分为 HIV-1 和 HIV-2 两个型。在分子生物学水平上，两者的核苷酸序列只有 45% 的同源性；在流行病学角度上，HIV-1 为全球性流行，且绝大部分艾滋病者是由 HIV-1 感染引起的，HIV-2 主要局限在非洲少数几个西非国家。

按照 1999 年 09 月发表的《HIV-1 的命名建议》[6]，现行的 HIV-1 的分类系统包括了四个单位：组（groups）、亚型（subtypes）、亚亚型（sub-subtypes）和流行重组型（CRFs）。根据编码包膜蛋白 env 基因和壳蛋白 gag 基因序列的同源性将 HIV-1 分为三个组[7]，M 组(main group)、O 组(outline group)和 N 组(non-M、non-O group)（图 1-2）。其中，M 组根据遗传差异性又可进一步分为 A~D、F~H、J、K 九种亚型和 14 种流行重组形式。在我国，以 HIV-1 为主要流行株，已发现的有 A、B（欧美 B）、B'（泰国 B）、C、D、E、F 和 G 8 个亚型及 2 个流行重组型（B'/C），以 B'、C、B'/C 重组和 E 亚型为主[8]。

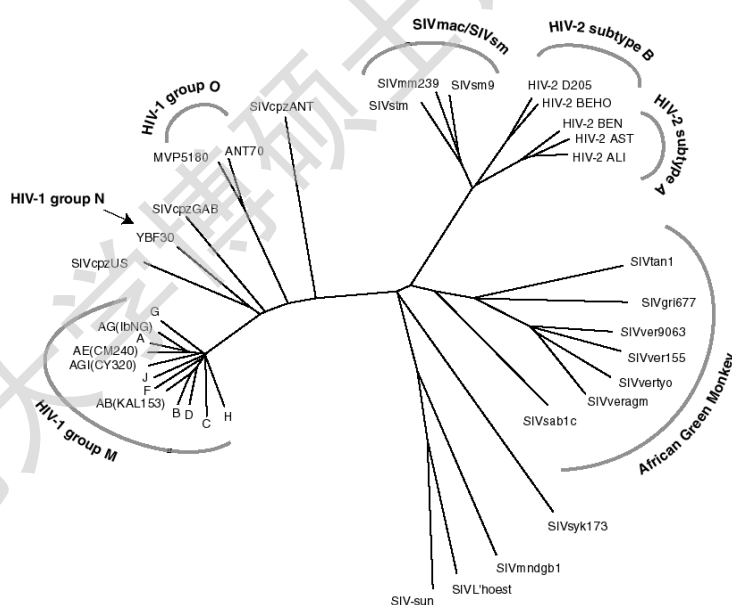


图 1-2 HIV-SIV 的进化树

Fig. 1-2 HIV-SIV phylogenetic tree

图片来源：HIV sequence database

1.2 HIV 的分子生物学特征

1.2.1 HIV 的颗粒结构

成熟的 HIV 病毒颗粒（图 1-3）在电镜下呈球形，类似 20 面体结构，直径

约 90~150 nm (图 1-3)。HIV 病毒具有慢病毒的特征, 即 p24 衣壳蛋白组成锥形核心。在此衣壳内含有两条相同的 RNA 链, 链上紧密结合着病毒的逆转录酶 RT (p66 和 p51)、整合酶和核衣壳蛋白 (p9 和 p6)。与病毒核心紧密连接的是 Vif 和 Nef 蛋白, 最近估计每个病毒粒子中含有 7~20 个 Vif 蛋白。在病毒颗粒内部 (很可能在核心外) 发现存在着病毒的辅助基因产物—Vpr。

病毒膜的内部由十四酰化的 p17 核心 (MA) 蛋白围绕[9]。包膜位于 HIV 病毒颗粒的最外层, 其主体为磷脂双分子层, 厚度约 7.5nm。同宿主细胞膜相比, HIV 的包膜在组成和功能上都存在明显的差异, 其中胆固醇同磷脂的摩尔比约为宿主细胞膜的 2.5 倍。包膜中所含的大量胆固醇, 可以消除磷脂的相变, 降低膜的通透性, 有利于保持病毒颗粒的稳定[10]。

包膜表面分布有 72 个刺突状结构, 该结构由糖蛋白 gp120 和 gp41 形成的三聚体或者四聚体组成。其中, gp41 为穿膜蛋白, 其跨膜结构域插入磷脂双分子层而固着在病毒的包膜上, 成为突起的“柄”, 通过非共价将 gp120 连接于包膜表面, 并参与包膜与宿主细胞膜的融合过程。有报道显示, gp41 的外部结构对中和抗体敏感[11]; gp120 则构成突起的头部, 与细胞膜上的特异受体相互作用, 从而介导病毒对宿主细胞的吸附, gp120 含有主要的中和表位[12]。

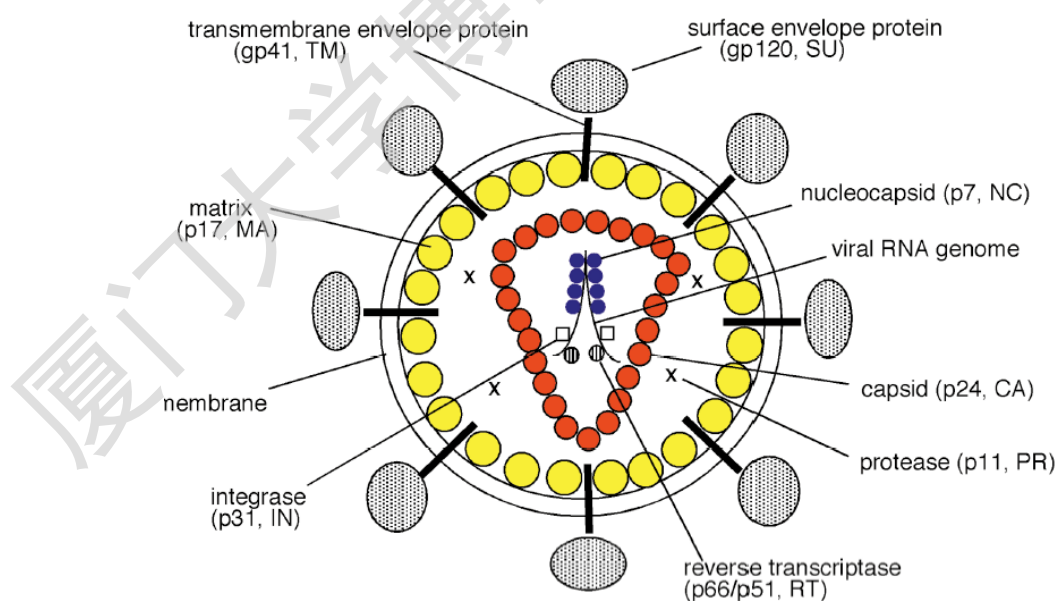


图 1-3 HIV-1 病毒颗粒结构[13]

Fig. 1-3 Schematic representation of the HIV-1 virion organization

1.2.2 HIV 的基因组结构与功能

HIV-1 的基因组全长约 9.8kb, 编码 3 种结构基因, 分别为 gag (group specific antigen) 基因、pol (polymerase) 基因和 env (envelope) 基因, 其中 gag 基因与 pol 基因使用不同的阅读框架, 并有部分重叠。另外还有编码 6 种调节蛋白的基因, 包括 HIV 复制必须的 tat (transactivator) 基因、rev (regulator of expression of virion) 基因以及另外 4 个编码辅助蛋白 (accessory protein) 的 vpr (viral protein R) 基因、vpu (viral protein U) 基因、vif (viral infectivity factor) 基因和 nef (negative factor) 基因[14] (图 1-4)。

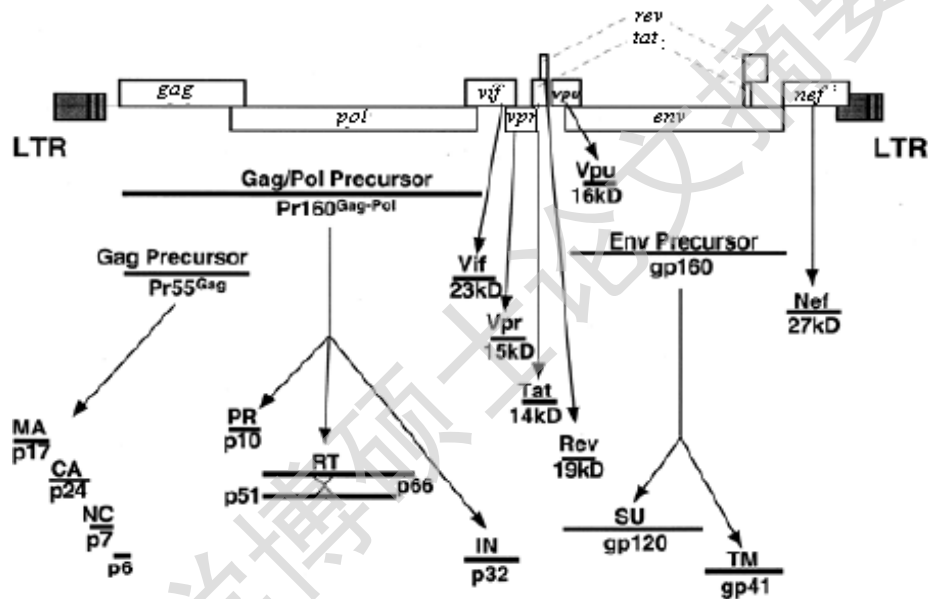


图 1-4 HIV-1 基因组[15]

Fig.1-4 Schematic of the HIV-1 genome

HIV-1 的基因组 RNA 中至少存在着 4 个剪接供体位点和 6 个剪接受体位点, 目前已发现 30 多种病毒 mRNA 转录加工后的产物。HIV 的基因可根据其表达时间的不同分为早期基因和晚期基因。在病毒表达的早期阶段, 病毒基因组转录出的 RNA 被剪接成分子量相对较低的 mRNA, 仅编码 Tat、Rev 和 Nef 等调节蛋白; 而病毒的晚期基因则包括核心蛋白、酶和包膜蛋白等, 分别由单剪接 mRNA 或全长的非剪接 mRNA 翻译而成, 在细胞中这些基因产物积累到一定数量后将进入病毒的组装阶段。

1.2.3 HIV 的生活周期

HIV 病毒是彻底的寄生生物，自身没有独立的代谢和能量转化系统，必须依在宿主细胞内才能表现出它的基本生命活力。HIV 病毒颗粒从亲代到子代的复制过程，构成了它的生活周期，包括：吸附、侵入脱壳、反转录、整合、病毒 RNA 和蛋白质合成、装配、释放和成熟等阶段（图 1-5）。

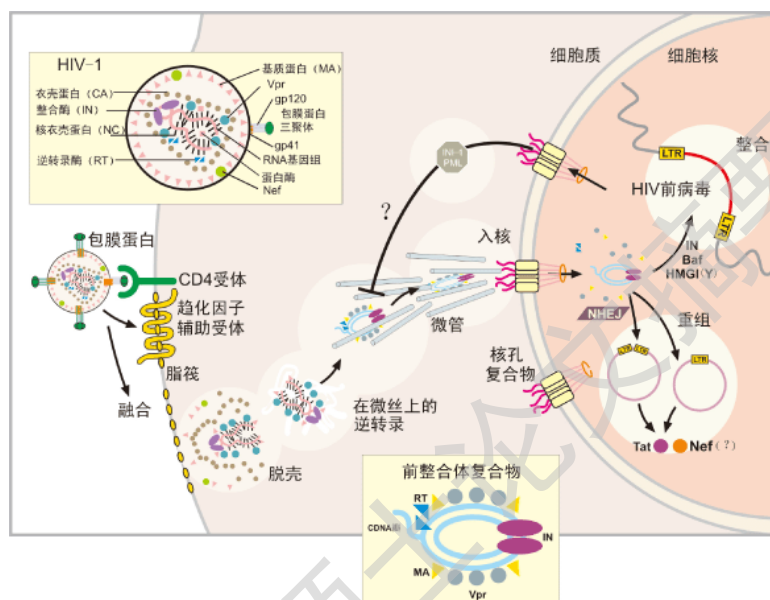


图 1-5 HIV 的生活史

Fig.1-5 Pathogenesis of HIV

图片来源: <http://www.hivinsite.ucsf.edu>

HIV 通过表面糖蛋白 gp120 和膜蛋白 gp41 多聚体与 CD4 分子结合，在趋化因子受体分子 CXCR4（T 细胞第二受体）或 CCR5（嗜单核-巨噬细胞的 HIV-1 第二受体）帮助下，病毒外膜与宿主细胞膜融合。进入宿主细胞内的 HIV 核心迅速脱壳，通过 p24 的去组装使衣壳解体。而后反转录出 HIV 双链 cDNA，利用其末端的 LTR 插入到宿主染色体 DNA 中，从而形成原病毒，可以随着宿主细胞的 DNA 复制而复制，造成长期潜伏感染。在静止的 CD4⁺细胞内原病毒不繁殖，一旦被激活，原病毒便利用宿主细胞的 RNA 聚合酶 II 转录形成全长病毒 mRNA，进入病毒的复制周期，表达结构蛋白及其它调节蛋白，再经过组装、芽生从细胞中释放，形成新的病毒，继续感染其它 CD4⁺细胞。

1.3 HIV 包膜糖蛋白的研究进展

HIV 包膜糖蛋白由 env 基因编码，env 基因长约 2.6kb。它先翻译出一个相

对分子量为 8.8×10^4 Da 的前体蛋白，经过糖基化加工后形成 HIV 包膜糖蛋白前体 gp160。gp160 由 845-870 个氨基酸残基组成，在去除了信号肽后（约 33-37 个氨基酸残基），于第 511 位残基处被宿主细胞所编码的蛋白酶裂解为表面糖蛋白（SU）gp120 和跨膜糖蛋白（TM）gp41 两部分（图 1-6），然后非共价结合成寡聚体（一般为三聚体），形成病毒包膜上的突起，介导病毒对宿主细胞的吸附和穿入。病毒表面糖蛋白的抗原性强于核心蛋白，可以从近 100% 的艾滋病患者血清中检测到抗 gp120 或 gp160 的抗体，高于 p24 抗体的阳性检出率[16]。

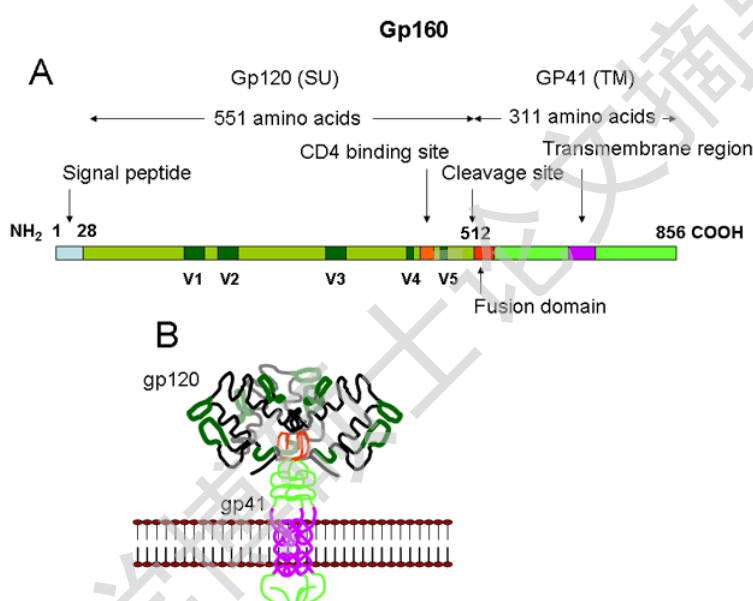


图 1-6 HIV 包膜糖蛋白 gp160 的结构 (A. gp160 结构域; B. 与病毒细胞膜相结合的 gp120/gp41 的三聚体)

Fig.1-6 The structure of gp160(A. The linear domain structure of gp160; B. A trimer of gp120/gp140 is associated with the viral membrane.)

图片来源: <http://pathmicro.med.sc.edu/lecture/hiv9.htm>

1.3.1 外膜糖蛋白 gp120

1.3.1.1 gp120 的结构特征

gp120 是一高度糖基化而又亲水的蛋白质，位于病毒颗粒或受感染细胞的表面，由 481 个氨基酸残基组成，相对分子量约 1.2×10^5 Da，至少含有 22 个糖基化位点，除 N-糖基化外，S 和 T 等氨基酸残基还可发生 O-糖基化。

在 HIV 编码的所有蛋白质中，gp120 的变异程度最高，易逃避机体免疫监视[17]。根据对 env 基因及 gp120 氨基酸序列的分析，可将其划分为 5 个高度变异

的区域（可变区，V1-V5）。这些可变区多暴露在 gp120 分子的表面，并被保守区（C1-C5）所分隔（图 1-7a）。

gp120 的晶体结构 X 射线衍射发现，其核心结构包括 25 个 β 链、5 个 α 螺旋和 10 个环状结构。gp120 的核心主要由保守区折叠而成，被一组 β 折叠结构分割为内部结构域（inner domain）和外部结构域（outer domain），这组 β 折叠（又称桥状结构 bridging sheet）由 V1 和 V2 环的茎部及 C4 区共 4 段反向平行的 β 链组成。由于缺少疏水核心，构象易发生改变。gp120 的内部结构域缺乏糖基化修饰，可能是寡聚化位点所在的部位，并与 N 端和 C 端肽段一起以非共价方式结合 gp41。外部结构则暴露于 gp120 三聚体的表面[17]（图 1-7 b,c）。

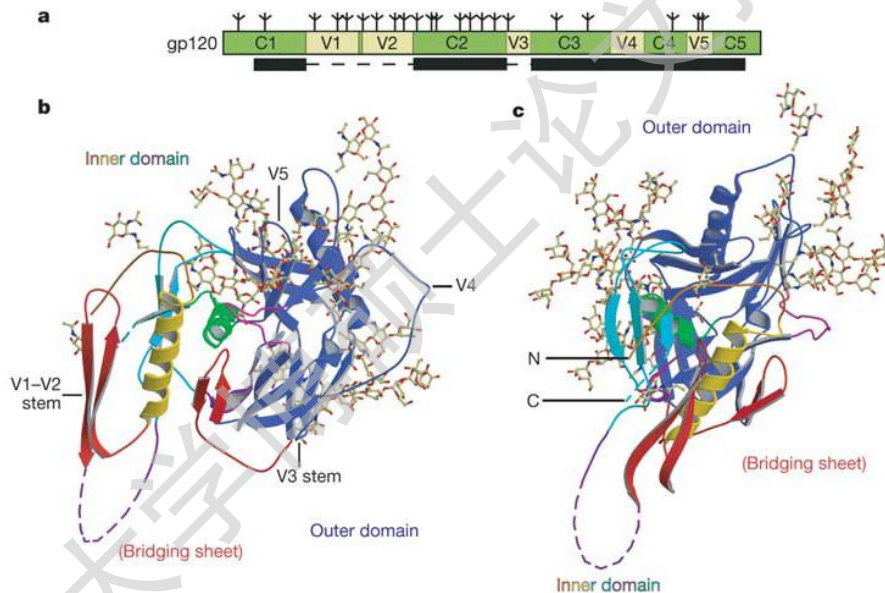


图 1-7 gp120 的结构[19] (a, gp120 的序列分析; b, 未结合的 gp120 结构;c, b 图 90 度旋转后)

Fig.1-7 The structure of gp120(a ,Diagram of sequence elements in gp120;b,Overall structure of the unliganded gp120;c,The same as b,rotated by 90° about a vertical axis)

1.3.1.2 gp120 介导病毒对宿主细胞的吸附

1.3.1.2.1 CD4（cluster of differentiation 4）

CD4 抗原是最早得到鉴定的反转录病毒受体，作为最主要的 HIV 受体蛋白，可与 HIV 表面蛋白 gp120 结合，介导病毒的吸附和侵入。CD4 分布于一部分 T 淋巴细胞、胸腺细胞和单核-巨噬细胞的表面，某些 B 淋巴细胞和脑细胞也表达

Degree papers are in the "[Xiamen University Electronic Theses and Dissertations Database](#)". Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.

厦门大学博硕士论文摘要库